

VP40\_EBV  
ID VP40\_EBV Reviewed; 605 AA.  
AC P03234;  
DT 21-JUL-1986, integrated into UniProtKB/Swiss-Prot.  
DT 21-JUL-1986, sequence version 1.  
DT 24-JUL-2007, entry version 64.  
DE Capsid protein P40 (Virion structural protein BVRF2) (EC-RF3 and EC-  
DE RF3A) [Contains: Capsid protein VP24 (Assemblin) (Protease)  
DE (EC 3.4.21.97); Capsid protein VP22A; C-terminal peptide].  
GN ORFNames=BVRF2;  
OS Epstein-Barr virus (strain B95-8) (HHV-4) (Human herpesvirus 4).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Gammaherpesvirinae; Lymphocryptovirus.  
OX NCBI\_TAXID=10377;  
OH NCBI\_TAXID=9606; Homo sapiens (Human).  
RN [1]  
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].  
RX MEDLINE=85035713; PubMed=6092825;  
RA Bankier A.T., Deininger P.L., Farrell P.J., Barrell B.G.;  
RT "Sequence analysis of the 17,166 base-pair EcoRI fragment C of B95-8  
RT Epstein-Barr virus.";  
RL Mol. Biol. Med. 1:21-45(1983).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RX MEDLINE=84270667; PubMed=6087149; DOI=10.1038/310207a0;  
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,  
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,  
RA Tuffnell P.S., Barrell B.G.;  
RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";  
RL Nature 310:207-211(1984).  
CC -!- FUNCTION: VP22A is a component of the capsid core involved in  
CC processing and packaging of progeny DNA. VP24 is a protease which  
CC can proteolytically cleave itself and VP22A at the C-terminus (By  
CC similarity).  
CC -!- CATALYTIC ACTIVITY: Cleaves -Ala-|-Ser- and -Ala-|-Ala- bonds in  
CC the scaffold protein.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative initiation; Named isoforms=2;  
CC Name=EC-RF3;  
CC IsoId=P03234-1; Sequence=Displayed;  
CC Name=EC-RF3A;  
CC IsoId=P03234-2; Sequence=VSP\_018865;  
CC -!- SIMILARITY: Belongs to the peptidase S21 family.  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL; V01555; CAA24801.1; -; Genomic\_DNA.  
DR PIR; A03798; QQBE3R.  
DR PDB; 106E; X-ray; A/B=1-235.  
DR MEROPS; S21.003; -.  
DR InterPro; IPR001847; Peptidase\_S21.  
DR Gene3D; G3DSA:3.20.16.10; Peptidase\_S21; 1.  
DR Pfam; PF00716; Peptidase\_S21; 1.  
DR PRINTS; PR00236; HSVCAPSIDP40.  
PE 1: Evidence at protein level;  
KW 3D-structure; Alternative initiation; Capsid protein; DNA packaging;

KW Hydrolase; Protease; Serine protease; Virion.

FT CHAIN 1 605 Capsid protein P40.  
/FTId=PRO\_0000027279.

FT CHAIN 1 235 Capsid protein VP24.  
/FTId=PRO\_0000027280.

FT CHAIN 236 ?568 Capsid protein VP22A.  
/FTId=PRO\_0000027281.

FT CHAIN ?569 605 C-terminal peptide.  
/FTId=PRO\_0000027283.

FT ACT\_SITE 48 48 Charge relay system (By similarity).  
FT ACT\_SITE 116 116 Charge relay system (By similarity).  
FT ACT\_SITE 139 139 Charge relay system (By similarity).  
FT SITE 235 236 Cleavage; by the protease (Probable).  
FT SITE 568 569 Cleavage; by the protease (Probable).  
FT VAR\_SEQ 1 260 Missing (in isoform EC-RF3A).  
/FTId=VSP\_018865.

FT STRAND 7 14  
FT STRAND 21 23  
FT HELIX 29 35  
FT STRAND 42 46  
FT STRAND 49 62  
FT STRAND 67 73  
FT HELIX 76 79  
FT HELIX 81 86  
FT HELIX 90 93  
FT HELIX 103 111  
FT STRAND 114 124  
FT STRAND 136 145  
FT STRAND 147 150  
FT STRAND 154 157  
FT HELIX 159 164  
FT HELIX 171 184  
FT HELIX 198 207  
FT HELIX 214 224  
FT STRAND 229 232

SQ SEQUENCE 605 AA; 64102 MW; FC2D355F8A389708 CRC64;

### Alignment Scores:

Pred. No.: 2e-88 Length: 605  
Score: 1894.00 Matches: 345  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 98.2% Indels: 0  
DB: 1 Gaps: 0

US-10-036-729-3 (1-1038) x VP40 EBV (1-605)

Qy	1	ATGCTATCAGGTAACGCAGGAGAAGGAGCACAGCCTGCGGAGGTTGGCCGCCGGGC 60
Db	261	MetLeuSerGlyAsnAlaGlyGluGlyAlaThrAlaCysGlyGlySerAlaAlaAlaGly 280
Qy	61	CAGGACCTCATCAGCGTCCCCGCAACACCTTATGACACTGCTTCAGACCAACCTGGAC 120
Db	281	GlnAspLeuIleSerValProArgAsnThrPheMetThrLeuLeuGlnThrAsnLeuAsp 300
Qy	121	AACAAACCGCCGAGGCAGACCCCGTACCCCTACGCGGCCCGCTGCCCTTCCCAC 180

Db 301 AsnLysProProArgGlnThrProLeuProTyrAlaAlaProLeuProProPheSerHis 320  
 Qy 181 CAGGCAATAGCCACCGCGCCTTCCTACGGTCTGGGCGGAGCGGTGCCCCGGCCGGC 240  
 |||||||  
 Db 321 GlnAlaIleAlaThrAlaProSerTyrGlyProGlyAlaGlyAlaValAlaProAlaGly 340  
 Qy 241 GGCTACTTTACCTCCCCCAGGAGGTTACTACGCCGGGCCGCGGGGGACCCGGGTGCC 300  
 |||||||  
 Db 341 GlyTyrPheThrSerProGlyGlyTyrTyrAlaGlyProAlaGlyAspProGlyAla 360  
 Qy 301 TTCTTGGCGATGGACGCTCACACCTACCACCCCCACCCACACCCCCCTCCGGCCTACTTT 360  
 |||||||  
 Db 361 PheLeuAlaMetAspAlaHisThrTyrHisProHisProProAlaTyrPhe 380  
 Qy 361 GGCTTGCAGGGCCTCTTGGCCCCCTCCACCCGTGCCTCCTACTACGGATCCCAC TTG 420  
 |||||||  
 Db 381 GlyLeuProGlyLeuPheGlyProProProProValProProTyrTyrGlySerHisLeu 400  
 Qy 421 CGGGCAGACTACGTCCCCGCTCCCTCGCGATCCAACAAGCGGAAAAGAGACCCCGAGGAG 480  
 |||||||  
 Db 401 ArgAlaAspTyrValProAlaProSerArgSerAsnLysArgLysArgAspProGluGlu 420  
 Qy 481 GATGAAGAAGGCAGGGGGCTATTCCGGGGGAGGACGCCACCCCTTACCGCAAGGACATA 540  
 |||||||  
 Db 421 AspGluGluGlyGlyLeuPheProGlyGluAspAlaThrLeuTyrArgLysAspIle 440  
 Qy 541 GCGGGCCTCTCAAGAGTGTGAATGAGTTACAGCACACGCTACAGGCCCTGCGCCGGAG 600  
 |||||||  
 Db 441 AlaGlyLeuSerLysSerValAsnGluLeuGlnHisThrLeuGlnAlaLeuArgArgGlu 460  
 Qy 601 ACGCTGTCCTACGGCCACACCGGAGTCGGATACTGCCCGCAGCAGGGCCCTGCTACACC 660  
 |||||||  
 Db 461 ThrLeuSerTyrGlyHisThrGlyValGlyTyrCysProGlnGlnGlyProCysTyrThr 480  
 Qy 661 CACTCGGGCCTACGGATTTCAGCCTCATCAAAGCTACGAAGTGCCAGATACTCCCT 720  
 |||||||  
 Db 481 HisSerGlyProTyrGlyPheGlnProHisGlnSerTyrGluValProArgTyrValPro 500  
 Qy 721 CATCCGCCCCCACCACCAACTCTCACCAAGGAGCTCAGGCGCAGGCCCTCCACCCCGGGC 780  
 |||||||  
 Db 501 HisProProProProThrSerHisGlnAlaAlaGlnAlaGlnProProProGly 520  
 Qy 781 ACACAGGCCCGAAGCCACTGTGTGGCCGAGTCCACGATCCCTGAGGCGGGAGCAGCC 840  
 |||||||  
 Db 521 ThrGlnAlaProGluAlaHisCysValAlaGluSerThrIleProGluAlaGlyAlaAla 540  
 Qy 841 GGGAACTCTGGACCCCGGGAGGACACCAACCCCTCAGCAGGCCACCCACCGAGGGCCACCAC 900  
 |||||||  
 Db 541 GlyAsnSerGlyProArgGluAspThrAsnProGlnGlnProThrThrGluGlyHisHis 560  
 Qy 901 CGCGGAAAGAAACTGGTGCAGGCCCTGCGTCCGGAGTGGCTCAGTCTAAGGAGCCCACC 960  
 |||||||  
 Db 561 ArgGlyLysLysLeuValGlnAlaSerAlaSerGlyValAlaGlnSerLysGluProThr 580  
 Qy 961 ACCCCCAAGGCCAAGTCTGTGTCAGCCCACCTCAAGTCCATCTTGCGAGGAATTGCTG 1020  
 |||||||  
 Db 581 ThrProLysAlaLysSerValSerAlaHisLeuLysSerIlePheCysGluGluLeuLeu 600

Qy 1021 AATAAACGCGTGGCT 1035  
| | | | | | | | | | | |  
Db 601 AsnLysArgValAla 605